PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

C2

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47, A61K 38/17, G01N 33/68, C12Q 1/68, C07K 16/18

A2

(11) International Publication Number: WO 99/61614

(43) International Publication Date: 2 December 1999 (02.12.99)

(21) International Application Number: PCT/US99/11497

(22) International Filing Date: 25 May 1999 (25.05.99)

(30) Priority Data:

60/087,104	28 May 1998 (28.05.98)	US
09/216,006	17 December 1998 (17.12.98)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/087,104 (CIP)
Filed on	28 May 1998 (28.05.98)
US	09/216,006 (CIP)
Filed on	17 December 1998 (17.12.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). GORGONE, Gina [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94547 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN SOCS PROTEINS

(57) Abstract

The invention provides human SOCS proteins (HSCOP) and polynucleotides which identify and encode HSCOP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HSCOP.

HUMAN SOCS PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human SOCS proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cancer, immune and neurological disorders. and infectious diseases.

BACKGROUND OF THE INVENTION

Signal transduction is a general process in which cells respond to extracellular signals (hormones, neurotransmitters, growth and differentiation factors, etc.) through a cascade of biochemical reactions beginning with the binding of the signal molecule to a cell membrane receptor and ending with an effect on an intracellular target molecule. Intermediate steps in this process involve the activation of various cytoplasmic proteins by phosphorylation via protein kinases and the translocation of some of these activated proteins to the cell nucleus, where the transcription of specific genes is affected. The signal transduction process regulates all types of cell functions, including cell proliferation, differentiation, and gene transcription.

Cytokines are a specific class of extracellular signaling molecules that control growth, differentiation, and various functions of hemopoietic and immune cells. Cytokines include interleukins (ILs), colony-stimulating factors (G-CSF and GM-CFS), erythropoietin (EPO), and various growth factors (EGF, PDGF, TGF, and FGF; Callard, R. and Gearing, A. (1994) The Cytokine Facts Book, pp 2-6, Academic Press, San Diego, CA).

Many of the cytokine receptors, including those for the growth factors EGF, PDGF, and FGF exhibit intrinsic protein kinase activity. Binding of the cytokine to its receptor triggers the autophosphorylation of a tyrosine residue on the receptor. It is believed that these phosphorylated residues are recognition sites for the binding of other cytoplasmic signaling proteins which link the initial receptor activation at the cell surface to the activation of a specific intracellular target molecule. These signaling proteins

and a distinctive motif N-terminal of the SOCS box. In addition to four new SOCS proteins containing the SH2 domain, three additional classes of SOCS proteins were found containing WD-40 repeats (WSB-1 and -2), SPRY domains (SSB-1 to -3), or ankyrin repeats (ASB-1 to -3). A class of small GTPases (Rar proteins) that contain the SOCS box were also identified. The function of WSB, SSB, and ASB proteins are as yet unknown. However, like SH2 domains, WD-40 repeats, ankyrin repeats, and SPRY domains have been implicated in protein-protein interactions (Hilton et al. supra).

Defects or alterations in the activity of signaling proteins such as CIS may play a role in the development of various proliferative disorders and diseases such as cancer. Loss or rearrangement of the putative human gene encoding CIS is associated with the development of renal cell carcinomas and lung cancer (Yoshimura et al., supra). This association suggests that CIS may function as a tumor suppressor gene.

The discovery of new human SOCS proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cancer, immune and neurological disorders, and infectious diseases.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human SOCS proteins, referred to collectively as "HSCOP" and individually as "HSOCP-1", "HSOCP-2", and "HSOCP-3", HSOCP-4", HSOCP-5", HSOCP-6", HSOCP-7", HSOCP-8", and HSOCP-9". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-9, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-9, and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HSCOP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods and algorithms used for identification of HSCOP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence

porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term “agonist” refers to a molecule which, when bound to HSCOP, increases or prolongs the duration of the effect of HSCOP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HSCOP.

An “allelic variant” is an alternative form of the gene encoding HSCOP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HSCOP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HSCOP or a polypeptide with at least one functional characteristic of HSCOP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HSCOP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HSCOP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HSCOP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HSCOP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HSCOP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HSCOP or fragments of HSCOP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and

conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with

nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HSCOP, or fragments thereof, or HSCOP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization

minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to HSCOP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human SOCS proteins (HSCOP), the polynucleotides encoding HSCOP, and the use of these compositions for the diagnosis, treatment, or prevention of cancer, immune and neurological disorders, and infectious diseases.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HSCOP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HSCOP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The clones in

of SEQ ID NO:18 from about nucleotide 973 to about nucleotide 1017. Polypeptides encoded by these fragments are useful, for example, as immunogenic peptides.

The invention also encompasses HSCOP variants. A preferred HSCOP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the HSCOP amino acid sequence, and which contains at least one functional or structural characteristic of HSCOP.

The invention also encompasses polynucleotides which encode HSCOP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:10-18, which encodes HSCOP.

The invention also encompasses a variant of a polynucleotide sequence encoding HSCOP. In particular, such a variant polynucleotide sequence will have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HSCOP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:10-18 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:10-18. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HSCOP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HSCOP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HSCOP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HSCOP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring

Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the Hamilton MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800

been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HSCOP may be cloned in recombinant DNA molecules that direct expression of HSCOP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HSCOP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HSOCH-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HSCOP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res.

expression vectors containing sequences encoding HSCOP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HSCOP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus,TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HSCOP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HSCOP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HSCOP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HSCOP are needed, e.g. for the production of antibodies, vectors which direct high level expression of HSCOP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HSCOP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris.

of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* or *apr⁻* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HSCOP is inserted within a marker gene sequence, transformed cells containing sequences encoding HSCOP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HSCOP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the

inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HSCOP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HSCOP may be designed to contain signal sequences which direct secretion of HSCOP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Bethesda MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HSCOP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HSCOP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HSCOP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using

to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as akathesia, Alzheimer's disease, amnesia, amyotrophic lateral sclerosis, bipolar disorder, catatonia, cerebral neoplasms, dementia, depression, diabetic neuropathy, Down's syndrome, tardive dyskinesia, dystonias, epilepsy, Huntington's disease, peripheral neuropathy, multiple sclerosis, neurofibromatosis, Parkinson's disease, paranoid psychoses, postherpetic neuralgia, schizophrenia, and Tourette's disorder; and an infectious disease such a viral infection, e.g., those caused by adenoviruses (acute respiratory disease, pneumonia), arenaviruses (lymphocytic choriomeningitis), bunyaviruses (Hantavirus), coronaviruses (pneumonia, chronic bronchitis), hepadnaviruses (hepatitis), herpesviruses (herpes simplex virus, varicella-zoster virus, Epstein-Barr virus, cytomegalovirus), flaviviruses (yellow fever), orthomyxoviruses (influenza), papillomaviruses (cancer), paramyxoviruses (measles, mumps), picornoviruses (rhinovirus, poliovirus, coxsackie-virus), polyomaviruses (BK virus, JC virus), poxviruses (smallpox), reovirus (Colorado tick fever), retroviruses

dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HSCOP may be produced using methods which are generally known in the art. In particular, purified HSCOP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HSCOP. Antibodies to HSCOP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HSCOP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HSCOP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HSCOP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HSCOP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984)

radioimmunoassay techniques may be used to assess the affinity of antibodies for HSCOP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of HSOCH-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HSCOP epitopes, represents the average affinity, or avidity, of the antibodies for HSCOP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HSCOP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the HSOCH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HSCOP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies. Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of HSOCH-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HSCOP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HSCOP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HSCOP. Thus, complementary molecules or fragments may be used to modulate HSCOP activity, or to achieve regulation of gene function. Such technology is now well known in the art,

hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HSCOP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HSCOP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino

example HSCOP or fragments thereof, antibodies of HSCOP, and agonists, antagonists or inhibitors of HSCOP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HSCOP may be used for the diagnosis of disorders characterized by expression of HSCOP, or in assays to monitor patients being treated with HSCOP or agonists, antagonists, or inhibitors of

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HSCOP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:10-18 or from genomic sequences including promoters, enhancers, and introns of the HSCOP gene.

Means for producing specific hybridization probes for DNAs encoding HSCOP include the cloning of polynucleotide sequences encoding HSCOP or HSCOP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HSCOP may be used for the diagnosis of disorders associated with expression of HSCOP. Examples of such disorders include, but are not limited to, a cancer such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis,

indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HSCOP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HSCOP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HSCOP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HSCOP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HSCOP, and will be employed under optimized conditions for identification of a specific gene or condition.

chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HSCOP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HSCOP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HSCOP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted

RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the REAL Prep 96 plasmid kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Str. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:10-18. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HSOCH-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HSCOP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HSCOP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HSOCH-encoding

metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified HSCOP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HSCOP Activity

HSCOP activity is demonstrated by the inhibition of differentiation in murine M1 cells transfected with the gene expressing HSCOP and induced to differentiate by treatment with IL-6 (Starr et al., supra). Differentiation is measured in the parent M1 cell line and in M1 cells transfected with HSCOP by the appearance of differentiated colonies arising from cells grown in semi-soft agar culture. The percent inhibition of differentiation in M1 transfected cells compared to the parent M1 cell line is proportional to the activity of HSCOP in the former cells.

XI. Functional Assays

HSCOP function is assessed by expressing the sequences encoding HSCOP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in

blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HSCOP Using Specific Antib dies

Naturally occurring or recombinant HSCOP is substantially purified by immunoaffinity chromatography using antibodies specific for HSCOP. An immunoaffinity column is constructed by covalently coupling anti-HSOCH antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HSCOP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HSCOP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HSOCH binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HSCOP is collected.

XIV. Identification of Molecules Which Interact with HSCOP

HSCOP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HSCOP, washed, and any wells with labeled HSCOP complex are assayed. Data obtained using different concentrations of HSCOP are used to calculate values for the number, affinity, and association of HSCOP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1 cont.

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
7	16	1275743	TESTTUT02	841412R1 (PROSTUT05), 1275743H1 (TESTTUT02), 2571403R6 (HIPOAZT01), 3319512F6 (PROSBPT03), 3964671X314D1 (PROSNOT14)
8	17	1722533	BLADNOT06	585432X15 (PROSNOT02), 1271329T6 (TESTTUT02), 1579164F1 (DUODNOT01), 1722533F6 and 1722533H1 (BLADNOT06), 2635492F6 (BONTNOT01), 2880628H1 and 2882159F6 (UTRSTUT05), 3203865H1 (PENCNOT02), 4852494H1 (TESTNOT10)
9	18	1759763	PITUNOT03	034803H1 (THP1NOB01), 161745R6 (ADENINB01), 595902H1 (BRAVUNT02), 626174R6 (PGANNOT01), 953742R1 (SCORNON01), 1759763H1 (PITUNOT03), 2138314F6 (ENDCNOT01), 2532454T6 (GBLANOT02), 3053743H1 (LNODNOT08)

Table 2 cont.

Protein Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Homologous Sequence	Analytical Methods
6	278	S219 T254 Y277		Ankyrin repeats: S52-D79 G85-N112 T117-E144 F150-N177 H182-Y209 SOCS box: P239-N278	Ankyrin protein	BLAST HMM PFAM
7	281	S33 T151 S166 S191 T97 T217 S256 S262	N272	Ras family domain: K16-C278 prenyl group binding site: C278-S281 SOCS box: V189-H237	Rar protein or Ras-like GTPase	BLAST, BLOCKS PRINTS, MOTIFS, PFAM, HMM
8	635	S413 S24 S26 S38 T47 T169 S247 T333 S545 S579 S6 T57 T180 S201 S345 T360 S390 S522 Y210	N202 N304 N331	Ankyrin repeats: E137-R364 S368-A400 R410-A472 SOCS box: P593-Q635	Ankyrin protein	BLAST PFAM HMM
9	518	T108 T109 S154 S234 S412 T270 S443 S473 S476	N106 N139	Ankyrin repeats: D9-M74 E78-S143 C145-N243 N279-C311 GTP-binding: L300-S317 SOCS box: V459-M506	Ankyrin protein	BLAST PFAM HMM BLOCKS

Table 3 cont.

Polynucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease, Disorder or Condition (Fraction of Total)	Vector
16	Reproductive (0.400) Nervous (0.360) Cardiovascular (0.080)	Cancer (0.680) Inflammation (0.120) Neurological (0.120)	pINCY
17	Reproductive (0.438) Gastrointestinal (0.156) Cardiovascular (0.094)	Cancer (0.438) Inflammation (0.250) Cell proliferation (0.156)	pINCY
18	Nervous (0.286) Cardiovascular (0.163) Hematopoietic/Immune (0.143)	Cancer (0.327) Cell proliferation (0.265) Inflammation (0.224)	pSPORT1

Table 4 cont.

Polynucleotide SEQ ID NO:	Library	Library Description
14	UTRSNOR01	The library was constructed using RNA isolated from nontumorous uterine endometrium tissue removed from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the endometrium was secretory, and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the associated tumor tissue indicated intramural uterine leiomyoma. Patient history included hypothyroidism and pelvic floor relaxation. Family history included benign hypertension, Type II diabetes, and hyperlipidemia.
15	KIDNNOT31	The library was constructed using RNA isolated from tissue that had kidney markers.
16	TESTTUT02	The library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.
17	BLADNOT06	The library was constructed using RNA isolated from the posterior wall bladder tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy and urinary diversion. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder and urothelium. Patient history included lung neoplasm, and tobacco abuse in remission. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
18	PITUNOT03	The library was constructed using RNA isolated from pituitary tissue of a 46-year-old Caucasian male who died from colon cancer.

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <u>supra</u>; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of HSCOP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
LAL, Preeti
HILLMAN, Jennifer L.
GORGONE, Gina
CORLEY, Neil C.
PATTERSON, Chandra
YUE, Henry
TANG, Y. Tom
AZIMZAI, Yalda

<120> HUMAN SOCS PROTEINS

<130> PF-0525 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/087,104; 09/216,006
<151> 1998-05-28; 1998-12-17

<160> 18

<170> PERL Program

<210> 1
<211> 288
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1758450

<400> 1
Met Ser Ser Ser Met Trp Tyr Ile Met Gln Ser Ile Gln Ser Lys
1 5 10 15
Tyr Ser Leu Ser Glu Arg Leu Ile Arg Thr Ile Ala Ala Ile Arg
20 25 30
Ser Phe Pro His Asp Asn Val Glu Asp Leu Ile Arg Gly Gly Ala
35 40 45
Asp Val Asn Cys Thr His Gly Thr Leu Lys Pro Leu His Cys Ala
50 55 60
Cys Met Val Ser Asp Ala Asp Cys Val Glu Leu Leu Leu Glu Lys
65 70 75
Gly Ala Glu Val Asn Ala Leu Asp Gly Tyr Asn Arg Thr Ala Leu
80 85 90
His Tyr Ala Ala Glu Lys Asp Glu Ala Cys Val Glu Val Leu Leu
95 100 105
Glu Tyr Gly Ala Asn Pro Asn Ala Leu Asp Gly Asn Arg Asp Thr
110 115 120
Pro Leu His Trp Ala Ala Phe Lys Asn Asn Ala Glu Cys Val Arg
125 130 135
Ala Leu Leu Glu Ser Gly Ala Ser Val Asn Ala Leu Asp Tyr Asn
140 145 150

Asn Asp Thr Pro Leu Ser Trp Ala Ala Met Lys Gly Asn Leu Glu
155 160 165
Ser Val Ser Ile Leu Leu Asp Tyr Gly Ala Glu Val Arg Val Ile
170 175 180
Asn Leu Ile Gly Gln Thr Pro Ile Ser Arg Leu Val Ala Leu Leu
185 190 195
Val Arg Gly Leu Gly Thr Glu Lys Glu Asp Ser Cys Phe Glu Leu
200 205 210
Leu His Arg Ala Val Gly His Phe Glu Leu Arg Lys Asn Gly Thr
215 220 225
Met Pro Arg Glu Val Ala Arg Asp Pro Gln Leu Cys Glu Lys Leu
230 235 240
Thr Val Leu Cys Ser Ala Pro Gly Thr Leu Lys Thr Leu Ala Arg
245 250 255
Tyr Ala Val Arg Arg Ser Leu Gly Leu Gln Tyr Leu Pro Asp Ala
260 265 270
Val Lys Gly Leu Pro Leu Pro Ala Ser Leu Lys Glu Tyr Leu Leu
275 280 285
Leu Leu Glu

<210> 2
<211> 423
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1834242

<400> 2
Met Lys Leu Thr Pro Arg Thr Ala Gly Arg Ala Trp Ala Gln Ser
1 5 10 15
Arg Lys Gly Lys Arg Ser Ser Trp Gly Gly Thr Ala Ala Val Ala
20 25 30
Glu Leu Lys Pro Gly Arg Pro His Gln Phe Asp Trp Lys Ser Ser
35 40 45
Cys Glu Thr Trp Ser Val Ala Phe Ser Pro Asp Gly Ser Trp Phe
50 55 60
Ala Trp Ser Gln Gly His Cys Ile Val Lys Leu Ile Pro Trp Pro
65 70 75
Leu Glu Glu Gln Phe Ile Pro Lys Gly Phe Glu Ala Lys Ser Arg
80 85 90
Ser Ser Lys Asn Glu Thr Lys Gly Arg Gly Ser Pro Lys Glu Lys
95 100 105
Thr Leu Asp Cys Gly Gln Ile Val Trp Gly Leu Ala Phe Ser Pro
110 115 120
Trp Pro Ser Pro Pro Ser Arg Lys Leu Trp Ala Arg His His Pro
125 130 135
Gln Val Pro Asp Val Ser Cys Leu Val Leu Ala Thr Gly Leu Asn
140 145 150
Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu Leu Leu
155 160 165
Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser Phe
170 175 180

Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys
185 190 195
Thr Leu Arg Ile Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln
200 205 210
Val Leu Ser Gly His Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser
215 220 225
Pro Asp Cys Ser Met Leu Cys Ser Ala Ala Gly Glu Lys Ser Val
230 235 240
Phe Leu Trp Ser Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu
245 250 255
Gly His Gln Ser Ser Val Val Ser Cys Asp Phe Ser Pro Asp Ser
260 265 270
Ala Leu Leu Val Thr Ala Ser Tyr Asp Thr Asn Val Ile Met Trp
275 280 285
Asp Pro Tyr Thr Gly Glu Arg Leu Arg Ser Leu His His Thr Gln
290 295 300
Val Asp Pro Ala Met Asp Asp Ser Asp Val His Ile Ser Ser Leu
305 310 315
Arg Ser Val Cys Phe Ser Pro Glu Gly Leu Tyr Leu Ala Thr Val
320 325 330
Ala Asp Asp Arg Leu Leu Arg Ile Trp Ala Leu Glu Leu Lys Thr
335 340 345
Pro Ile Ala Phe Ala Pro Met Thr Asn Gly Leu Cys Cys Thr Phe
350 355 360
Phe Pro His Gly Gly Val Ile Ala Thr Gly Thr Arg Asp Gly His
365 370 375
Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser Leu Lys His
380 385 390
Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr Gln Val
395 400 405
Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr Tyr
410 415 420
Arg Thr Phe

<210> 3
<211> 349
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1849725

<400> 3
Met Glu Asp Pro Gln Ser Lys Glu Pro Ala Gly Glu Ala Val Ala
1 5 10 15
Pro Ala Leu Leu Glu Ser Pro Arg Pro Glu Gly Gly Glu Glu Pro
20 25 30
Pro Arg Pro Ser Pro Glu Glu Thr Gln Gln Cys Lys Phe Asp Gly
35 40 45
Gln Glu Thr Lys Gly Ser Lys Phe Ile Thr Ser Ser Ala Ser Asp
50 55 60
Phe Ser Asp Pro Val Tyr Lys Glu Ile Ala Ile Thr Asn Gly Cys
65 70 75
Ile Asn Arg Met Ser Lys Glu Glu Leu Arg Ala Lys Leu Ser Glu

```
                        80                      85                      90
Phe Lys Leu Glu Thr Arg Gly Val Lys Asp Val Leu Lys Lys Arg
                        95                      100                     105
Leu Lys Asn Tyr Tyr Lys Lys Gln Lys Leu Met Leu Lys Glu Ser
                        110                     115                     120
Asn Phe Ala Asp Ser Tyr Tyr Asp Tyr Ile Cys Ile Ile Asp Phe
                        125                     130                     135
Glu Ala Thr Cys Glu Glu Gly Asn Pro Pro Glu Phe Val His Glu
                        140                     145                     150
Ile Ile Glu Phe Pro Val Val Leu Leu Asn Thr His Thr Leu Glu
                        155                     160                     165
Ile Glu Asp Thr Phe Gln Gln Tyr Val Arg Pro Glu Ile Asn Thr
                        170                     175                     180
Gln Leu Ser Asp Phe Cys Ile Ser Leu Thr Gly Ile Thr Gln Asp
                        185                     190                     195
Gln Val Asp Arg Ala Asp Thr Phe Pro Gln Val Leu Lys Lys Val
                        200                     205                     210
Ile Asp Trp Met Lys Leu Lys Glu Leu Gly Thr Lys Tyr Lys Tyr
                        215                     220                     225
Ser Leu Leu Thr Asp Gly Ser Trp Asp Met Ser Lys Phe Leu Asn
                        230                     235                     240
Ile Gln Cys Gln Leu Ser Arg Leu Lys Tyr Pro Pro Phe Ala Lys
                        245                     250                     255
Lys Trp Ile Asn Ile Arg Lys Ser Tyr Gly Asn Phe Tyr Lys Val
                        260                     265                     270
Pro Arg Ser Gln Thr Lys Leu Thr Ile Met Leu Glu Lys Leu Gly
                        275                     280                     285
Met Asp Tyr Asp Gly Arg Pro His Cys Gly Leu Asp Asp Ser Lys
                        290                     295                     300
Asn Ile Ala Arg Ile Ala Val Arg Met Leu Gln Asp Gly Cys Glu
                        305                     310                     315
Leu Arg Ile Asn Glu Lys Met His Ala Gly Gln Leu Met Ser Val
                        320                     325                     330
Ser Ser Ser Leu Pro Ile Glu Gly Thr Pro Pro Pro Gln Met Pro
                        335                     340                     345
His Phe Arg Lys



<210> 4
<211> 355
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 2547840

<400> 4
Met Ala Arg Arg Pro Arg Asn Ser Arg Ala Trp His Phe Val Leu
 1              5                   10                  15
Ser Ala Ala Arg Arg Asp Ala Asp Ala Arg Ala Val Ala Leu Ala
                20                  25                  30
Gly Ser Thr Asn Trp Gly Tyr Asp Ser Asp Gly Gln His Ser Asp
                35                  40                  45
Ser Asp Ser Asp Pro Glu Tyr Ser Thr Leu Pro Pro Ser Ile Pro
```

```
                    50                       55                       60
Ser Ala Val Pro Val Thr Gly Glu Ser Phe Cys Asp Cys Ala Gly
                    65                       70                       75
Gln Ser Glu Ala Ser Phe Cys Ser Ser Leu His Ser Ala His Arg
                    80                       85                       90
Gly Arg Asp Cys Arg Cys Gly Glu Glu Asp Glu Tyr Phe Asp Trp
                    95                      100                      105
Val Trp Asp Asp Leu Asn Lys Ser Ser Ala Thr Leu Leu Ser Cys
                   110                      115                      120
Asp Asn Arg Lys Val Ser Phe His Met Glu Tyr Ser Cys Gly Thr
                   125                      130                      135
Ala Ala Ile Arg Gly Thr Lys Glu Leu Gly Glu Gly Gln His Phe
                   140                      145                      150
Trp Glu Ile Lys Met Thr Ser Pro Val Tyr Gly Thr Asp Met Met
                   155                      160                      165
Val Gly Ile Gly Thr Ser Asp Val Asp Leu Asp Lys Tyr Arg His
                   170                      175                      180
Thr Phe Cys Ser Leu Leu Gly Arg Asp Glu Asp Ser Trp Gly Leu
                   185                      190                      195
Ser Tyr Thr Gly Leu Leu His His Lys Gly Asp Lys Thr Ser Phe
                   200                      205                      210
Ser Ser Arg Phe Gly Gln Gly Ser Ile Ile Gly Val His Leu Asp
                   215                      220                      225
Thr Trp His Gly Thr Leu Thr Phe Phe Lys Asn Arg Lys Cys Ile
                   230                      235                      240
Gly Val Ala Ala Thr Lys Leu Gln Asn Lys Arg Phe Tyr Pro Met
                   245                      250                      255
Val Cys Ser Thr Ala Ala Arg Ser Ser Met Lys Val Thr Arg Ser
                   260                      265                      270
Cys Ala Ser Ala Thr Ser Leu Gln Tyr Leu Cys Cys His Arg Leu
                   275                      280                      285
Arg Gln Leu Arg Pro Asp Ser Gly Asp Thr Leu Glu Gly Leu Pro
                   290                      295                      300
Leu Pro Pro Gly Leu Lys Gln Val Leu His Asn Lys Leu Gly Trp
                   305                      310                      315
Val Leu Ser Met Ser Cys Ser Arg Arg Lys Ala Pro Val Ser Asp
                   320                      325                      330
Pro Gln Ala Ala Thr Ser Ala His Pro Ser Ser Arg Glu Pro Arg
                   335                      340                      345
Pro Cys Gln Arg Lys Arg Cys Arg Arg Thr
                   350                      355
```

<210> 5
<211> 421
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 3071986

<400> 5

```
Met Ala Ser Phe Pro Pro Arg Val Asn Glu Lys Glu Ile Val Arg
  1               5                  10                  15
```

```
Leu Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala Ala Pro Phe Asp
                20                  25                  30
Lys Lys Cys Gly Arg Glu Asn Trp Thr Val Ala Phe Ala Pro Asp
                35                  40                  45
Gly Ser Tyr Phe Ala Trp Ser Gln Gly His Arg Thr Val Lys Leu
                50                  55                  60
Val Pro Trp Ser Gln Cys Leu Gln Asn Phe Leu Leu His Gly Thr
                65                  70                  75
Lys Asn Val Thr Asn Ser Ser Ser Leu Arg Leu Pro Arg Gln Asn
                80                  85                  90
Ser Asp Gly Gly Gln Lys Asn Lys Pro Arg Glu His Ile Ile Asp
                95                  100                 105
Cys Gly Asp Ile Val Trp Ser Leu Ala Phe Gly Ser Ser Val Pro
                110                 115                 120
Glu Lys Gln Ser Arg Cys Val Asn Ile Glu Trp His Arg Phe Arg
                125                 130                 135
Phe Gly Gln Asp Gln Leu Leu Leu Ala Thr Gly Leu Asn Asn Gly
                140                 145                 150
Arg Ile Lys Ile Trp Asp Val Tyr Thr Gly Lys Leu Leu Leu Asn
                155                 160                 165
Leu Val Asp His Thr Glu Val Val Arg Asp Leu Thr Phe Ala Pro
                170                 175                 180
Asp Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys Thr Leu
                185                 190                 195
Arg Val Trp Asp Leu Lys Asp Asp Gly Asn Met Met Lys Val Leu
                200                 205                 210
Arg Gly His Gln Asn Trp Val Tyr Ser Cys Ala Phe Ser Pro Asp
                215                 220                 225
Ser Ser Met Leu Cys Ser Val Gly Ala Ser Lys Ala Val Phe Leu
                230                 235                 240
Trp Asn Met Asp Lys Tyr Thr Met Ile Arg Lys Leu Glu Gly His
                245                 250                 255
His His Asp Val Val Ala Cys Asp Phe Ser Pro Asp Gly Ala Leu
                260                 265                 270
Leu Ala Thr Ala Ser Tyr Asp Thr Arg Val Tyr Ile Trp Asp Pro
                275                 280                 285
His Asn Gly Asp Ile Leu Met Glu Phe Gly His Leu Phe Pro Pro
                290                 295                 300
Pro Thr Pro Ile Phe Ala Gly Gly Ala Asn Asp Arg Trp Val Arg
                305                 310                 315
Ser Val Ser Phe Ser His Asp Gly Leu His Val Ala Ser Leu Ala
                320                 325                 330
Asp Asp Lys Met Val Arg Phe Trp Arg Ile Asp Glu Asp Tyr Pro
                335                 340                 345
Val Gln Val Ala Pro Leu Ser Asn Gly Leu Cys Cys Ala Phe Ser
                350                 355                 360
Thr Asp Gly Ser Val Leu Ala Ala Gly Thr His Asp Gly Ser Val
                365                 370                 375
Tyr Phe Trp Ala Thr Pro Arg Gln Val Pro Ser Leu Gln His Leu
                380                 385                 390
Cys Arg Met Ser Ile Arg Arg Val Met Pro Thr Gln Glu Val Gln
                395                 400                 405
Glu Leu Pro Ile Pro Ser Lys Leu Leu Glu Phe Leu Ser Tyr Arg
                410                 415                 420
Ile
```

<210> 6
<211> 278
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 3484619

<400> 6
Met Glu Pro Arg Ala Ala Asp Gly Cys Phe Leu Gly Asp Val Gly
1 5 10 15
Phe Trp Val Glu Arg Thr Pro Val His Glu Ala Ala Gln Arg Gly
20 25 30
Glu Ser Leu Gln Leu Gln Gln Leu Ile Glu Ser Gly Ala Cys Val
35 40 45
Asn Gln Val Thr Val Asp Ser Ile Thr Pro Leu His Ala Ala Ser
50 55 60
Leu Gln Gly Gln Ala Arg Cys Val Gln Leu Leu Leu Ala Ala Gly
65 70 75
Ala Gln Val Asp Ala Arg Asn Ile Asp Gly Ser Thr Pro Leu Cys
80 85 90
Asp Ala Cys Ala Ser Gly Ser Ile Glu Cys Val Lys Leu Leu Leu
95 100 105
Ser Tyr Gly Ala Lys Val Asn Pro Pro Leu Tyr Thr Ala Ser Pro
110 115 120
Leu His Glu Ala Cys Met Ser Gly Ser Ser Glu Cys Val Arg Leu
125 130 135
Leu Ile Asp Val Gly Ala Asn Leu Glu Ala His Asp Cys His Phe
140 145 150
Gly Thr Pro Leu His Val Ala Cys Ala Arg Glu His Leu Asp Cys
155 160 165
Val Lys Val Leu Leu Asn Ala Gly Ala Asn Val Asn Ala Ala Lys
170 175 180
Leu His Glu Thr Ala Leu His His Ala Ala Lys Val Lys Asn Val
185 190 195
Asp Leu Ile Glu Met Leu Ile Glu Phe Gly Gly Asn Ile Tyr Ala
200 205 210
Arg Asp Asn Arg Gly Lys Lys Pro Ser Asp Tyr Thr Trp Ser Ser
215 220 225
Ser Ala Pro Ala Lys Cys Phe Glu Tyr Tyr Glu Lys Thr Pro Leu
230 235 240
Thr Leu Ser Gln Leu Cys Arg Val Asn Leu Arg Lys Ala Thr Gly
245 250 255
Val Arg Gly Leu Glu Lys Ile Ala Lys Leu Asn Ile Pro Pro Arg
260 265 270
Leu Ile Asp Tyr Leu Ser Tyr Asn
275

<210> 7
<211> 281
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1275743

<400> 7
Met Gly Ser Gln Gly Ser Pro Val Lys Ser Tyr Asp Tyr Leu Leu
1 5 10 15
Lys Phe Leu Leu Val Gly Asp Ser Asp Val Gly Lys Gly Glu Ile
20 25 30
Leu Glu Ser Leu Gln Asp Gly Ala Ala Glu Ser Pro Tyr Ala Tyr
35 40 45
Ser Asn Gly Ile Asp Tyr Lys Thr Thr Thr Ile Leu Leu Asp Gly
50 55 60
Arg Arg Val Lys Leu Glu Leu Trp Asp Thr Ser Gly Gln Gly Arg
65 70 75
Phe Cys Thr Ile Phe Arg Ser Tyr Ser Arg Gly Ala Gln Gly Ile
80 85 90
Leu Leu Val Tyr Asp Ile Thr Asn Arg Trp Ser Phe Asp Gly Ile
95 100 105
Asp Arg Trp Ile Lys Glu Ile Asp Glu His Ala Pro Gly Val Pro
110 115 120
Arg Ile Leu Val Gly Asn Arg Leu His Leu Ala Phe Lys Arg Gln
125 130 135
Val Pro Thr Glu Gln Ala Arg Ala Tyr Ala Glu Lys Asn Cys Met
140 145 150
Thr Phe Phe Glu Val Ser Pro Leu Cys Asn Phe Asn Val Ile Glu
155 160 165
Ser Phe Thr Glu Leu Ser Arg Ile Val Leu Met Arg His Gly Met
170 175 180
Glu Lys Ile Trp Arg Pro Asn Arg Val Phe Ser Leu Gln Asp Leu
185 190 195
Cys Cys Arg Ala Ile Val Ser Cys Thr Pro Val His Leu Ile Asp
200 205 210
Lys Leu Pro Leu Pro Val Thr Ile Lys Ser His Leu Lys Ser Phe
215 220 225
Ser Met Ala Asn Gly Met Asn Ala Val Met Met His Gly Arg Ser
230 235 240
Tyr Ser Leu Ala Ser Gly Ala Gly Gly Gly Gly Ser Lys Gly Asn
245 250 255
Ser Leu Lys Arg Ser Lys Ser Ile Arg Pro Pro Gln Ser Pro Pro
260 265 270
Gln Asn Cys Ser Arg Ser Asn Cys Lys Ile Ser
275 280

<210> 8
<211> 635
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1722533

<400> 8
Met Ala Thr Gln Ile Ser Thr Arg Gly Ser Gln Cys Thr Ile Gly
1 5 10 15

Gln Glu Glu Tyr Ser Leu Tyr Ser Ser Leu Ser Glu Asp Glu Leu
20 25 30
Val Gln Met Ala Ile Glu Gln Ser Leu Ala Asp Lys Thr Arg Gly
35 40 45
Pro Thr Thr Ala Glu Ala Thr Ala Ser Ala Cys Thr Asn Arg Gln
50 55 60
Pro Ala His Phe Tyr Pro Trp Thr Arg Ser Thr Ala Pro Pro Glu
65 70 75
Ser Ser Pro Ala Arg Ala Pro Met Gly Leu Phe Gln Gly Val Met
80 85 90
Gln Lys Tyr Ser Ser Ser Leu Phe Lys Thr Ser Gln Leu Ala Pro
95 100 105
Ala Asp Pro Leu Ile Lys Ala Ile Lys Asp Gly Asp Glu Glu Ala
110 115 120
Leu Lys Thr Met Ile Lys Glu Gly Lys Asn Leu Ala Glu Pro Asn
125 130 135
Lys Glu Gly Trp Leu Pro Leu His Glu Ala Ala Tyr Tyr Gly Gln
140 145 150
Val Gly Cys Leu Lys Val Leu Gln Arg Ala Tyr Pro Gly Thr Ile
155 160 165
Asp Gln Arg Thr Leu Gln Glu Glu Thr Ala Val Tyr Leu Ala Thr
170 175 180
Cys Arg Gly His Leu Asp Cys Leu Leu Ser Leu Leu Gln Ala Gly
185 190 195
Ala Glu Pro Asp Ile Ser Asn Lys Ser Arg Glu Thr Pro Leu Tyr
200 205 210
Lys Ala Cys Glu Arg Lys Asn Ala Glu Ala Val Lys Ile Leu Val
215 220 225
Gln His Asn Ala Asp Thr Asn His Arg Cys Asn Arg Gly Trp Thr
230 235 240
Ala Leu His Glu Ser Val Ser Arg Asn Asp Leu Glu Val Met Gln
245 250 255
Ile Leu Val Ser Gly Gly Ala Lys Val Glu Ser Lys Asn Ala Tyr
260 265 270
Gly Ile Thr Pro Leu Phe Val Ala Ala Gln Ser Gly Gln Leu Glu
275 280 285
Ala Leu Arg Phe Leu Ala Lys Tyr Gly Ala Asp Ile Asn Thr Gln
290 295 300
Ala Ser Asp Asn Ala Ser Ala Leu Tyr Glu Ala Cys Lys Asn Glu
305 310 315
His Glu Glu Val Val Glu Phe Leu Leu Ser Gln Gly Ala Asp Ala
320 325 330
Asn Lys Thr Asn Lys Asp Gly Leu Leu Pro Leu His Ile Ala Ser
335 340 345
Lys Lys Gly Asn Tyr Arg Ile Val Gln Met Leu Leu Pro Val Thr
350 355 360
Ser Arg Thr Arg Ile Arg Arg Ser Gly Val Ser Pro Leu His Leu
365 370 375
Ala Ala Glu Arg Asn His Asp Glu Val Leu Glu Ala Leu Leu Ser
380 385 390
Ala Arg Phe Asp Val Asn Thr Pro Leu Ala Pro Glu Arg Ala Arg
395 400 405
Leu Tyr Glu Asp Arg Arg Thr Ser Ala Leu Tyr Phe Ala Val Val
410 415 420
Asn Asn Asn Val Tyr Ala Thr Glu Leu Leu Leu Gln His Gly Ala
425 430 435
Asp Pro Asn Arg Asp Val Ile Ser Pro Leu Leu Val Ala Ile Arg

```
                440                     445                     450
His Gly Cys Leu Arg Thr Met Gln Leu Leu Leu Asp His Gly Ala
                455                     460                     465
Asn Ile Asp Ala Tyr Ile Ala Thr His Pro Thr Ala Phe Pro Ala
                470                     475                     480
Thr Ile Met Phe Ala Met Lys Cys Leu Ser Leu Leu Lys Phe Leu
                485                     490                     495
Met Asp Leu Gly Cys Asp Gly Glu Pro Cys Phe Ser Cys Leu Tyr
                500                     505                     510
Gly Asn Gly Pro His Pro Pro Ala Pro Gln Pro Ser Ser Arg Phe
                515                     520                     525
Asn Asp Ala Pro Ala Ala Asp Lys Glu Pro Ser Val Val Gln Phe
                530                     535                     540
Cys Glu Phe Val Ser Ala Pro Glu Val Ser Arg Trp Ala Gly Pro
                545                     550                     555
Ile Ile Asp Val Leu Leu Asp Tyr Val Gly Asn Val Gln Leu Cys
                560                     565                     570
Ser Arg Leu Lys Glu His Ile Asp Ser Phe Glu Asp Trp Ala Val
                575                     580                     585
Ile Lys Glu Lys Ala Glu Pro Pro Arg Pro Leu Ala His Leu Cys
                590                     595                     600
Arg Leu Arg Val Arg Lys Ala Ile Gly Lys Tyr Arg Ile Lys Leu
                605                     610                     615
Leu Asp Thr Leu Pro Leu Pro Gly Arg Leu Ile Arg Tyr Leu Lys
                620                     625                     630
Tyr Glu Asn Thr Gln
                635
```

<210> 9
<211> 518
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1759763

<400> 9

```
Met Asp Phe Thr Glu Ala Tyr Ala Asp Thr Cys Ser Thr Val Gly
  1               5                  10                  15
Leu Ala Ala Arg Glu Gly Asn Val Lys Val Leu Arg Lys Leu Leu
                 20                  25                  30
Lys Lys Gly Arg Ser Val Asp Val Ala Asp Asn Arg Gly Trp Met
                 35                  40                  45
Pro Ile His Glu Ala Ala Tyr His Asn Ser Val Glu Cys Leu Gln
                 50                  55                  60
Met Leu Ile Asn Ala Asp Ser Ser Glu Asn Tyr Ile Lys Met Lys
                 65                  70                  75
Thr Phe Glu Gly Phe Cys Ala Leu His Leu Ala Ala Ser Gln Gly
                 80                  85                  90
His Trp Lys Ile Val Gln Ile Leu Leu Glu Ala Gly Ala Asp Pro
                 95                 100                 105
Asn Ala Thr Thr Leu Glu Glu Thr Thr Pro Leu Phe Leu Ala Val
                110                 115                 120
```

```
Glu Asn Gly Gln Ile Asp Val Leu Arg Leu Leu Leu Gln His Gly
                125                 130                 135
Ala Asn Val Asn Gly Ser His Ser Met Cys Gly Trp Asn Ser Leu
                140                 145                 150
His Gln Ala Ser Phe Gln Glu Asn Ala Glu Ile Ile Lys Leu Leu
                155                 160                 165
Leu Arg Lys Gly Ala Asn Lys Glu Cys Gln Asp Asp Phe Gly Ile
                170                 175                 180
Thr Pro Leu Phe Val Ala Ala Gln Tyr Gly Lys Leu Glu Ser Leu
                185                 190                 195
Ser Ile Leu Ile Ser Ser Gly Ala Asn Val Asn Cys Gln Ala Leu
                200                 205                 210
Asp Lys Ala Thr Pro Leu Phe Ile Ala Ala Gln Glu Gly His Thr
                215                 220                 225
Lys Cys Val Glu Leu Leu Leu Ser Ser Gly Ala Asp Pro Asp Leu
                230                 235                 240
Tyr Cys Asn Glu Asp Ser Trp Gln Leu Pro Ile His Ala Ala Ala
                245                 250                 255
Gln Met Gly His Thr Lys Ile Leu Asp Leu Leu Ile Pro Leu Thr
                260                 265                 270
Asn Arg Ala Cys Asp Thr Gly Leu Asn Lys Val Ser Pro Val Tyr
                275                 280                 285
Ser Ala Val Phe Gly Gly His Glu Asp Cys Leu Glu Ile Leu Leu
                290                 295                 300
Arg Asn Gly Tyr Ser Pro Asp Ala Gln Ala Cys Leu Val Phe Gly
                305                 310                 315
Phe Ser Ser Pro Val Cys Met Ala Phe Gln Lys Asp Cys Glu Phe
                320                 325                 330
Phe Gly Ile Val Asn Ile Leu Leu Lys Tyr Gly Ala Gln Ile Asn
                335                 340                 345
Glu Leu His Leu Ala Tyr Cys Leu Lys Tyr Glu Lys Phe Ser Ile
                350                 355                 360
Phe Arg Tyr Phe Leu Arg Lys Gly Cys Ser Leu Gly Pro Trp Asn
                365                 370                 375
His Ile Tyr Glu Phe Val Asn His Ala Ile Lys Ala Gln Ala Lys
                380                 385                 390
Tyr Lys Glu Trp Leu Pro His Leu Leu Val Ala Gly Phe Asp Pro
                395                 400                 405
Leu Ile Leu Leu Cys Asn Ser Trp Ile Asp Ser Val Ser Ile Asp
                410                 415                 420
Thr Leu Ile Phe Thr Leu Glu Phe Thr Asn Trp Lys Thr Leu Ala
                425                 430                 435
Pro Ala Val Glu Arg Met Leu Ser Ala Arg Ala Ser Asn Ala Trp
                440                 445                 450
Ile Leu Gln Gln His Ile Ala Thr Val Pro Ser Leu Thr His Leu
                455                 460                 465
Cys Arg Leu Glu Ile Arg Ser Ser Leu Lys Ser Glu Arg Leu Arg
                470                 475                 480
Ser Asp Ser Tyr Ile Ser Gln Leu Pro Leu Pro Arg Ser Leu His
                485                 490                 495
Asn Tyr Leu Leu Tyr Glu Asp Val Leu Arg Met Tyr Glu Val Pro
                500                 505                 510
Glu Leu Ala Ala Ile Gln Asp Gly
                515
```

<210> 10
<211> 1117
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1758450

<400> 10
```
cacgccttga cagcggcttt caaccccac ctcagcccag caattcggca gtttggagca   60
tgtgaacacc ttgagccttg atgagttcca gtatgtggta tattatgcag agcattcaga 120
gcaaatactc tctctccgag cgcttaatcc gaacaattgc tgccatccgt tccttcccac 180
atgataatgt agaggacctc atcagagggg gagcagatgt gaactgcact catggcacac 240
tgaagccctt gcactgtgcc tgtatggtgt cagatgctga ctgtgtggag ttacttctgg 300
aaaaaggagc cgaggtgaat gccctggatg ggtataaccg aacagccctc cactatgcag 360
cagagaaaga tgaggcttgt gtggaggtcc tattggagta tggtgcaaac cccaatgctt 420
tggatggcaa cagagatacc ccacttcact gggcagcctt taagaacaat gctgagtgtg 480
tgcgggctct cctagagagc ggggcctctg tcaatgccct ggattacaac aatgatacac 540
cgctcagctg ggctgccatg aagggaaatc ttgagagtgt cagcatcctt ctggattatg 600
gcgcagaggt cagagtcatc aacctaatag gccagacacc catctcccgc ctggtggctc 660
tgctagtcag gggacttgga acagagaaag aggactcttg ctttgagctc ctccacagag 720
ctgttggaca ctttgaattg aggaaaaatg gcaccatgcc acgagaggtg gccagagacc 780
cgcagctatg tgaaaaactg actgttctgt gctcagctcc aggaactcta aaaacactcg 840
ctcgctatgc cgtgcgccgt agcctgggac tccagtatct ccccgatgca gtgaagggcc 900
ttccactgcc agcttctttg aaggaatacc tgttactttt agaatagccg gagaagatgt 960
ttgcaccatc gtgcaggcag ctctgggtga ggttgtccct gcagtactcc ttgtcacaga 1020
aaacagaaaa acagttgttt cctgatgtgt gggttataga tttcgaagca acatgtcaca 1080
acaataacct gcatagcaac tccccttcc aaacaaa                          1117
```

<210> 11
<211> 2589
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1834242

<400> 11
```
cttgaatgaa gctgacacca agaaccgcgg gaagagcttg ggcccaaagc aggaaaggga   60
agcgctcgag ttggggagga accgctgctg tggccgaact caagcccggg cgccccccacc 120
agtttgattg gaagtccagc tgtgaaacct ggagcgtcgc cttctcccca gatggctcct 180
ggtttgcttg gtctcaagga cactgcatcg tcaaactgat cccctggccg ttggaggagc 240
agttcatccc taaagggttt gaagccaaaa gccgaagtag caaaaatgag acgaaagggc 300
ggggcagccc aaaagagaag acgctggact gtggtcagat tgtctggggg ctggccttca 360
gcccgtggcc ttccccaccc agcaggaagc tctgggcacg ccaccacccc caagtgcccg 420
atgtctcttg cctggttctt gctacgggac tcaacgatgg gcagatcaag atctgggagg 480
tgcagacagg gctcctgctt ttgaatcttt ccggccacca agatgtcgtg agagatctga 540
gcttcacacc cagtggcagt ttgattttgg tctccgcgtc acgggataag actcttcgca 600
tctgggacct gaataaacac ggtaaacaga ttcaagtgtt atcgggccac ctgcagtggg 660
tttactgctg ttccatctcc ccagactgca gcatgctgtg ctctgcagct ggagagaagt 720
cggtctttct atggagcatg aggtcctaca cgttaattcg gaagctagag ggccatcaaa 780
gcagtgttgt ctcttgtgac ttctcccccg actctgccct gcttgtcacg gcttcttacg 840
ataccaatgt gattatgtgg gacccctaca ccggcgaaag gctgaggtca ctccaccaca 900
cccaggttga ccccgccatg gatgacagtg acgtccacat tagctcactg agatctgtgt 960
```

gcttctctcc ggaaggcttg taccttgcca cggtggcaga tgacagactc ctcaggatct 1020
gggccctgga actgaaaact cccattgcat ttgctcctat gaccaatggg ctttgctgca 1080
catttttcc acatggtgga gtcattgcca cagggacaag agatggccac gtccagttct 1140
ggacagctcc tagggtcctg tcctcactga agcacttatg ccggaaagcc cttcgaagtt 1200
tcctaacaac ttaccaagtc ctagcactgc caatccccaa gaaaatgaaa gagttcctca 1260
catacaggac tttttaagca acaccacatc ttgtgcttct ttgtagcagg gtaaatcgtc 1320
ctgtcaaagg gagttgctgg aataatgggc caaacatctg gtcttgcatt gaaatagcat 1380
ttctttggga ttgtgaatag aatgtagcaa aaccagattc cagtgtacta gtcatggatc 1440
tttctctccc tggcatgtga aagtcagtct tagaggaaga gattccactt gcacggcaac 1500
agagccttac gttaaatctt cagtccagtt atgaacagca agtgttgaac tctttctgct 1560
tgttttgatt caaagtgcag ttactgatgt tgttttgatt atgcaactaa gtaggcctcc 1620
agagcctctc tagtggcaga gcagctcaca ctccctccgc tgggaacgat ggcttctgcc 1680
tagtacctat ccttgtgttt ctgatgcagt ggtagcattg gttcaagttc tctcctgctg 1740
tggtcagagt tgcttcgatg ttggccaagt gcttttcttc ttgggctccc ttctgacctg 1800
caggacagtt ttcctggagc catttggtat gaggtattaa tttagcttaa ctaaattaca 1860
ggggactcag aggccgtgct cctgaccgat ccagacacta ttactggctt tttttttttt 1920
tttttaacaa tggtgtgcat gtgcaggaaa tgacaaattt gtatgtcaga ttatacaagg 1980
atgtattctt aaaccgcatg actattcaga tggctactga gttatcagtg gccatttatt 2040
agcatcatat ttatttgtat tttctcaaca gatgttaagg tacaactgtg tttttctcga 2100
ttatctaaaa accatagtac ttaaattgaa cagttgcaaa gatgtcttaa ttgtgtaaag 2160
aattggtgta gtcatgactt tagctgatac tcttatgtac gagatctgtc tctgctgttt 2220
aacttcattg gattaatcag ctggtttcaa ctctactgcg aaacaaaaat agctccttaa 2280
aagtactgtt ctccttcagt ggcatgtagt tatctaatca agacacctca ttcaaacaaa 2340
acctgcctta ggaaaattta atatatttta aattatttta aaagaaatac aacatcttat 2400
tctttagctt tcttaatcgg tgctttatgg aggccagtgt aacgttacat gactcgttga 2460
gaaagttgag gaatttcctc taccaccttt gttgcttgaa gaaaaacatg tcttttcaaa 2520
atgagaggct ttcattgaag aaaagaaaaa aacaacagtt aaaagctaaa aaaaaaaaaa 2580
aaaaaaaaa 2589

<210> 12
<211> 2038
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1849725

<400> 12
cgggaacgcg agcccggtaa tttttcaacg gagaaaggcg aggctttcgg gctctgcaga 60
gtgagagtta gcaagtgtcc ggctccagca actctcctct ggcgtgacag ccggcatgga 120
ggatccacag agtaaagagc ctgccggcga ggccgtggct cccgcgctgc tggagtcgcc 180
gcggccggag ggcggggagg agccgccgcg tcccagtccc gaggaaactc aacagtgtaa 240
atttgatggc caggagacaa aaggatccaa gttcattacc tccagtgcga gtgacttcag 300
tgacccggtt tacaaagaga ttgccattac gaatggctgt attaatagaa tgagtaagga 360
agaactcaga gctaagcttt cagaattcaa gcttgaaact agaggagtaa aggatgttct 420
aaagaagaga ctgaaaaact attataagaa gcagaagctg atgctgaaag agagcaattt 480
tgctgacagt tattatgact acatttgtat tattgacttt gaagccactt gtgaagaagg 540
aaacccacct gagtttgtac atgaaataat tgaatttccg gttgttttac tgaatacgca 600
tactttagaa atagaagaca cgtttcagca gtatgtaaga ccagagatta acacacagct 660
gtctgatttc tgcatcagtc taactggaat tactcaggat caggtagaca gagctgatac 720
cttccctcag gtactaaaaa aagtaattga ctggatgaaa ttgaaggaat taggaacaaa 780
gtataaatac tcacttttaa cagatggttc ttgggatatg agtaagttct tgaacattca 840
gtgtcaactc agcaggctca aataccctcc ttttgcgaaa aagtggatca atattcggaa 900
gtcatatgga aatttttaca aggttcctag aagccaaacc aaactgacaa taatgcttga 960
aaaattagga atggattatg atgggcggcc tcactgtggt cttgatgact ctaagaatat 1020

cgcccgaata gcagttcgaa tgcttcagga tgggtgtgaa ctccgaatca acgagaaaat 1080
gcatgcagga cagctaatga gtgtgtcctc ttccttacca atagagggca ctccaccacc 1140
acaaatgcca cattttagaa agtaacagtt ttgtgtgtgg atcattccaa ttgaagttgc 1200
tatgaagagg tagcagatga atctcattga attagtcctg tagtgcaaac tttaagcacc 1260
ttaaaacatt taaaatctta ttacaggtga tagagataga tacatgtatg tgaacagatt 1320
ttgtaggaag gcatactgaa ttctttgtca ccaagcactt ttgataatgg acaggaatcc 1380
ggtaacctag ataaccaagg tcctggtcta acacaatggg atattttaat aattttaaag 1440
agggggttcc acaggttata aattcccttt ttttgggtgt ttaaaaaaat ggcccaaaaa 1500
tctcctaaat atggggcctt ggtgtctctc ggtttggaaa atgggccaac aaatcccttt 1560
aactcggggg tggtggttta acaaataaat gggtagaaat ggggtggggt tttccccttt 1620
taaatttaaa accatttcca ccttaaggga ttggtaaaca cccccettaa atcccettta 1680
aaaaaattgg tcccgggaaa aattgggatt tgggggcaaa agggtaagga attcctgtaa 1740
tccctaaagg cctctctttg ggggaatttt tccccagggg gaatatatcc ccttaaggtg 1800
ccccccttt gtggaatttt tttccccaaa aggggtttat aataaatgtt gggaaaagtt 1860
ttccaccccc aaggggaaat gggtgggggt gggaaaattt tccggtaaaa gaggtgacac 1920
tttggggtag atgacccata aatacttgcg cctcaagggg gtttgcccct attttcaaaa 1980
aactccccta aaaatttggg gggaggagaa ttttatttgg attagggggg tttatata 2038

<210> 13
<211> 1537
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 2547840

<400> 13
tgggtcggac gcgtgggcgg aaagtgggtc agggccgggc cggcggagcg cgcagcgggg 60
cagccagatt ctttccacca tggccagacg cccccggaac agcagggcct ggcacttcgt 120
cctgagtgca gcccgccgag acgcagatgc ccgggccgtg gctctagcag gctccactaa 180
ctggggctac gactctgatg ggcagcacag cgactcggac tccgaccccg agtactccac 240
gctgccgcca tccatcccca gtgcggtgcc cgtgaccggc gagtccttct gtgactgtgc 300
tgggcagagc gaggcctcct tctgtagcag cctgcactcg gcccaccggg gcagggactg 360
ccgctgcgga gaggaagacg agtatttcga ctgggtctgg gatgacttaa ataagtcatc 420
agccaccctg ctgagctgtg acaaccgtaa ggtcagcttc cacatggagt acagctgcgg 480
cacagcggcc atccggggca ccaaggagct gggggagggc cagcacttct gggagatcaa 540
gatgacctct cccgtctacg gcaccgacat gatggtgggc atcgggacgt cggatgtgga 600
cctggacaaa taccgccaca cgttctgcag cctgctgggc agggatgagg acagctgggg 660
cctctcctac acgggcctcc tccaccacaa gggcgacaag accagcttct cgtcgcggtt 720
cggccagggc tccatcattg gcgtgcacct ggacacctgg cacggcacac tcaccttttt 780
caagaacagg aagtgtatag gtgtggcagc caccaagctg cagaacaaga gattctaccc 840
gatggtgtgc tccacggcgg cccggagcag catgaaggtc acccgctcct gtgccagcgc 900
cacttccctc cagtacctgt gctgccaccg cctgcgccag ctgcggccag actcgggaga 960
cacgctggag ggtctgccgc tgccgccggg cctcaagcag gtgctacaca acaagctggg 1020
ctgggtcctg agcatgagtt gcagccgccg caaggctcca gtgtccgatc cccaggcagc 1080
gacctccgcc caccccagca gtcgcgagcc tcggccctgc cagaggaagc gctgccgccg 1140
gacctgactg acttcccagt ggaactgcct tcttgggctg ggacagcccc tttcctctgt 1200
cccttctttc tctgtccctt ccttccagcc acactccagg gcggagttgg atgaggcccg 1260
tccggaggga gccatctctt gctcccgagg ctgggacagt cctttctgtg gggctctag 1320
ggcccctctg ctgctgtgct gggtggggaa gcggctgccc tgagccccag gtcttgtggg 1380
aggctgcgag gacgagagcc tggctggagc ccgcgttgct gttcccacag ggcctcggtt 1440
tttcctaact tgctctgcat gctgtcagcg gctgccccgc cgtcatagac ttaaaggact 1500
gcaataaatg tagagttgat gtctaacaaa aaaaaaa 1537

<210> 14
<211> 2203
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 3071986

<400> 14
```
ctgtcttcct ccgcagcgcg aggctgggta cagggtctat tgtctgtggt tgactccgta   60
ctttggtctg aggccttcgg gagctttccc gaggcagtta gcagaagccg cagcggccgc  120
ccccgcccgt ttcctctgtc cctgggcccg ggagggacca acttggcgtc acgcccctca  180
gcggtcgcca ctctcttctc tgttgttggg tccgcatcgt attcccggaa tcagacggtg  240
ccccatagat ggccagcttt cccccgaggg tcaacgagaa agagatcgtg agattacgta  300
ctataggtga acttttagct cctgcagctc cttttgacaa gaaatgtggt cgtgaaaatt  360
ggactgttgc ttttgctcca gatggttcat actttgcttg gtcacaagga catcgcacag  420
taaagcttgt tccgtggtcc cagtgccttc agaactttct cttgcatggc accaagaatg  480
ttaccaattc aagcagttta agattgccaa gacaaaatag tgatggtggt cagaaaaata  540
agcctcgtga acatattata gactgtggag atatagtctg gagtcttgct tttgggtcat  600
cagttccaga aaaacagagt cgctgtgtaa atatagaatg gcatcgcttc agatttggac  660
aagatcagct acttcttgct acagggttga acaatgggcg tatcaaaata tgggatgtat  720
atacaggaaa actcctcctt aacttggtag atcatactga agtggtcaga gatttaactt  780
ttgctccaga tggaagcttg atcctggtgt cagcttcaag agacaaaact ctcagagtat  840
gggacctgaa agatgatgga aacatgatga aagtattgag gggcatcag aattgggtgt  900
acagctgtgc attctctcct gactcttcta tgctgtgttc agtcggagcc agtaaagcag  960
ttttcctttg gaatatggat aaatacacca tgatacggaa actagaagga catcaccatg 1020
atgtggtagc ttgtgacttt tctcctgatg gagcattact ggctactgca tcttatgata 1080
ctcgagtata tatctgggat ccacataatg gagacattct gatggaattt gggcacctgt 1140
ttcccccacc tactccaata tttgctggag gagcaaatga ccggtgggta cgatctgtat 1200
cttttagcca tgatggactg catgttgcaa gccttgctga tgataaaatg gtgaggttct 1260
ggagaattga tgaggattat ccagtgcaag ttgcaccttt gagcaatggt ctttgctgtg 1320
ccttctctac tgatggcagt gttttagctg ctgggacaca tgacggaagt gtgtattttt 1380
gggccactcc acggcaggtc cctagcctgc aacatttatg tcgcatgtca atccgaagag 1440
tgatgcccac ccaagaagtt caggagctgc cgattccttc caagcttttg gagtttctct 1500
cgtatcgtat ttagaagatt ctgccttccc tagtagtagg gactgacaga atacacttaa 1560
cacaaacctc aagctttact gacttcaatt atctgttttt aaagacgtag aagatttatt 1620
taatttgata tgttcttgta ctgcattttg atcagttgag cttttaaaat attatttata 1680
gacaatagaa gtatttctga acatatcaaa tataaatttt tttaaagatc taactgtgaa 1740
aacatacata cctgtacata tttagatata agctgctata tgttgaatgg acccttttgc 1800
ttttctgatt tttagttctg acatgtatat attgcttcag tagagccaca atatgtatct 1860
ttgctgtaaa gtgcaaggaa attttaaatt ctgggacact gagttagatg gtaaatactg 1920
acttacgaaa gttgaattgg gtgaggcggg caaatcacct gaggtcagca gtttgagact 1980
agcctggcaa acatgatgaa accctgtctc tactaaaaat acaaaaaaaa aaaaaattag 2040
ccaggcgtgg tggtgcacac ctgtagtcct agctacttgg gaggctgagg caggagaatt 2100
gcttgaaccc aggaggtgga ggttgcagta agccaagatc acaccactgc actccaacct 2160
ggacaacaga gcgagactcc atctcaaaaa aaaaataaaa agg                   2203
```

<210> 15
<211> 1622
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 3484619

<400> 15
ccgcatggag ccccgggcgg cggacggctg cttcctgggc gacgtgggtt tctgggtgga 60
gcggacccct gtgcacgagg cagcccagcg gggtgagagc ctgcagctgc aacagctgat 120
cgagagcggc gcctgcgtga accaggtcac cgtggactcc atcacgcccc tgcacgcagc 180
cagtctgcag ggccaggcgc ggtgtgtgca gctgctgctg gcggctgggg cccaggtgga 240
tgctcgcaac atcgacggca gcaccccgct ctgcgatgcc tgcgcctcgg gcagcatcga 300
gtgtgtgaag ctcttgctgt cctacggggc caaggtcaac cctcccctgt acacagcgtc 360
ccccctgcac gaggcctgca tgagcgggag ttccgaatgt gtgaggcttc ttattgacgt 420
cggggccaat ctggaagcgc acgattgcca ttttgggacc cctctgcacg ttgcctgtgc 480
ccgggagcat ctggactgtg tcaaagtgct gctcaatgca ggggccaacg tgaatgcggc 540
aaagcttcat gagactgccc ttcaccacgc ggccaaggtc aagaatgttg acctcatcga 600
gatgcttatc gagtttggcg gcaacatcta cgcccgggac aaccgcggga agaagccgtc 660
tgactacacg tggagcagca gcgctcccgc caagtgcttc gagtactacg aaaagacacc 720
tctgactctg tcacagctct gcagggtgaa cttgaggaag gccactggcg tccgagggct 780
ggagaagatt gccaagttaa acatcccgcc ccggctcatt gattacctct cctacaactg 840
aattgcaggt ggggtccgga ccgtgactgc ccccgttgtg cccagcattg cccgggtgag 900
ggctctgcct gttcctctga agcagcgtga ttgctgtaga tagaacaacg ctccttcgag 960
tcccttcctg cgatcctgtt taggcttctc tcctggatcc tggataatgt ttccagggtg 1020
ttgggaaggc ctgcgtctca ggtcacagtt gtgggtgtgg ccctgcgctg ttctacagaa 1080
cctaccctct caatgggcat gggcccaacc atccagtttt cctcttttac ggaccatcct 1140
caaaggcact ctcaggacag acggcgtggg gagcacagag gaggctggca gagctgggga 1200
ctgagggcat tgttgctgat tctcactcac cggggcagcc tgccgcagat gcacaggccc 1260
caggtgcagg ccaccacctc cgggtcggca ccaggactgc cctcggtgct catagggaat 1320
ggctgggccc acggaaggtc ggcctgggat gtggcctggg actgctgctc tgctggctgc 1380
tgtgtggatg cttttcctgg agcactttcc aaggcatccc ccagccccaa gcctgcgcgc 1440
atctgtcact cagggacttt ctatgggtct ttgtggggga aggccctggc tttgtattcc 1500
cacaagtagc actgagtttc ttaggaaatt tgtcttcagt attaagtctc caactcttgt 1560
aaaaagttta tatgtaggat aaaaaccttt tagaggacac gtaggcggta ccactaaggt 1620
tt 1622

<210> 16
<211> 1385
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1275743

<400> 16
cgcaacgggc gcaggtgcgg ggcgcgggct ctctcacgcc gcggcctcac ccggcggtgc 60
ttcggcaggc ggccggcgcg gggcgcaggc ggcgcggcca tgggctcgca gggcagtccg 120
gtgaagagct acgactacct gctcaagttc ctgctggtgg gcgacagcga cgtgggcaag 180
ggcgagatcc tggagagcct gcaggacggc gcggcagagt ccccgtacgc ctacagtaac 240
gggatcgact acaagaccac caccatcctg ctggacggcc ggcgcgtgaa gctggagctc 300
tgggacacgt cgggccaggg ccggttctgc accatcttca ggtcctactc cagggcgct 360
caggggatcc tcttggtgta tgacatcacc aaccgctggt cctttgacgg catcgaccgc 420
tggatcaagg agatcgatga gcatgcaccc ggagtccccc ggatcttggt tggaaaccgg 480
ctgcacctgg ccttcaagcg gcaggtcccg acggagcagg cccgcgcgta cgcagagaag 540
aactgcatga ccttctttga ggtcagcccc ctgtgcaact tcaacgtcat cgagtccttc 600
acggagctat cccgcatcgt gctcatgcgg cacggcatgg agaagatctg gaggcccaac 660
cgagtgttca gcctgcagga cctctgctgc cgggccatcg tctcctgcac ccccgtgcac 720
ctcatcgaca agcttccact gcccgtcacc atcaagagcc acctcaagtc cttctcgatg 780
gccaacggca tgaacgcggt catgatgcac ggccgttcct actccctggc cagcggggcc 840
gggggcggcg gcagcaaggg caacagcctc aagaggtcca agtccatccg tccaccccag 900
agcccccccc agaactgctc gcggagtaac tgcaagatct cctagcgggg atgggcgggg 960

ccgcctgtgc agatgccagg agggctcgag ctggacactc ctggctggac gccaggccag 1020
tgccgcctac gtggagactg tccacacagc tgcctcagaa gcgccgggct ttcctcacac 1080
ctgagccggg tgcgaggagg agcatgcacg gaccaagcgc ggcaggcggg aggaggggc 1140
gcggctgggc tgctggtgct tccgggaatc ttggtcggaa acaagccggg cctccccagc 1200
tgcctgggct tgaccggcgg ggagcctggt tggcctttct tatttatata gagaacactt 1260
cactttttg tacattttta aggggccttc agggaagcct gggtgtggcc cggtggtggt 1320
gcactggtga cttcatggcc acgccagctg cggggacgca cttgggactc ctcgagaggg 1380
gactc 1385

<210> 17
<211> 2790
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1722533

<400> 17
cccggaggaa cggaaggcag gattgcagct tcctcagtgc aacctccaaa caggaaatct 60
gagatgtata acatcttata gtttggcttg tcaacgttgg tcatgcggtg gccccaaaat 120
aaactccctg cttcaaagga cagcgtttca gaactgcctg gcagagcagc cagaagcttg 180
gggccagggc agaaggaaaa ctcggggagc atgttctgaa ttaagacact ttcaagaaaa 240
tcctttgtat taccccctgaa ttgtaccctt gtttcagagc ctaacagggt tttctgattt 300
gctgttccct cctccccact gggtgtctgt tctggaggcc agggtgagag gtggaggagg 360
atggccacgc agatcagcac tcggggcagc cagtgtacca ttgggcagga ggagtacagc 420
ctgtacagca gcctgagcga ggatgaactg gtgcagatgg ccatcgagca gagcctagcg 480
gacaagacaa ggggcccaac cactgctgag gccaccgcgt ctgcatgtac caaccgccaa 540
cctgcccatt tctacccatg gaccaggtcc actgcacctc ctgagagttc gccggcccgg 600
gccccaatgg gcttgttcca aggggtcatg cagaaataca gcagcagctt gttcaagacc 660
tcccagctgg cgcctgcgga cccccttgata aaggccatca aggatggcga tgaagaggcc 720
ttgaagacca tgatcaagga agggaagaat ctcgcagagc ccaacaagga gggctggctg 780
ccgctgcacg aggccgcata ctatggccag gtgggctgcc tgaaagtcct gcagcgagcg 840
tacccaggga ccatcgacca gcgcaccctg caggaggaaa cagccgttta cttggcaacg 900
tgcaggggcc acctggactg tctcctgtca ctgctccaag caggggcaga gccggacatc 960
tccaacaaat cccgagagac accgctctac aaagcctgtg agcgcaagaa cgcggaggcc 1020
gtgaagattc tggtgcagca caacgcagac accaaccacc gctgcaaccg cggctggacc 1080
gctctgcacg agtctgtgtc tcgcaatgac ctggaggtca tgcagatcct ggtgagcgga 1140
ggagccaagg tggaatccaa gaacgcctac ggcatcaccc ccttgttcgt ggccgcccag 1200
agtggacagt tggaggcctt gaggttctta gccaagtacg gtgctgacat caacacgcag 1260
gccagcgaca acgcgtctgc cctctacgag gcctgcaaga atgagcatga ggaggtggtg 1320
gagtttctgc tgtcacaggg tgccgacgcc aacaagacca acaaggacgg cttgctcccg 1380
ctgcacatcg cctccaagaa gggcaactac aggatcgtgc agatgctgct gccggtgacc 1440
agccgcacgc gcatacgccg tagcggcgtc agtccgctgc acctggcggc cgagcgcaac 1500
cacgacgagg tgctggaggc gctgctgagc gcgcgcttcg acgtgaacac gccgctggcg 1560
cccgaacgcg cgcgcctcta cgaagaccgg cgcacgtccg cgctgtactt cgcggtggtc 1620
aacaacaacg tgtacgccac cgagctgctg ctgcaacacg gcgccgaccc caaccgcgac 1680
gtcatcagcc ccttgctcgt ggccatccgc cacggctgcc tgcgcacaat gcagctgctg 1740
ctggaccacg gcgcgaacat cgacgcctat atcgccacgc accccaccgc cttccccgcc 1800
accatcatgt tcgccatgaa gtgcctgtcg ctgctcaagt tcctcatgga cctggcctgc 1860
gacggcgagc cctgcttctc atgcctctac ggcaacggcc cgcacccgcc ggccccgcag 1920
ccctccagca ggttcaacga cgcgcccgcg gccgacaagg agcccagcgt ggtgcagttc 1980
tgtgagttcg tatctgcccc agaggtgagc cgctgggcgg ggcccatcat cgatgtcctc 2040
ctggactacg tgggcaacgt gcagctctgc tcgcggctga aggaacacat cgacagcttt 2100

gaggactggg ccgtcatcaa ggagaaggca gaacctccaa gacctctggc tcacctttgc 2160
cgactgcggg ttcgaaaggc cattgggaaa taccgtataa aactcctaga caccttgccg 2220
ctcccaggca ggctgattag atacctgaaa tacgagaaca cccagtaact ggggccacgg 2280
ggagagagga gtagcccctc agactcttct tactaagtct caggacgtcg gtgttcccaa 2340
ctccaagggg acctggtgac agacgaggct gcaggctgcc tccctctcag cctggacagc 2400
taccaggatc tcactgggtc tcagggccca gagctttggc cagagcagag aacagaatgt 2460
gtcaaggaga agaatcattt gtttacaaac tgatgagcag atcccagacc ttctctacct 2520
tcaggaatgg cagaaacctc tattcctggg gccagggcag agcttgaggt gttctgggga 2580
aggtggtgct cagagccttc cctgtgcccc tccacttgtt ctggaaaact caccacttga 2640
cttcagagct ttctctccaa agactaagat gaagacgtgg cccaaggtag gggtaggggg 2700
gagcctgggt cttggagggc tttgttaagt attaatataa taaatgttac acatgtgaca 2760
cctgcccagt gaaaaactaa aaaaaaaaaa 2790

<210> 18
<211> 2263
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte clone 1759763

<400> 18
cggaccgtcc gcctacgggg gccgggacgt cgcctgcgcg tctctcgttt tcggacggct 60
gcagcatcgc ggtggggatc gaaagcgggg gcttctggga cgcagctctg gagacgcggc 120
ctcggaccag ccatttcggt gtagaagtgg cagcacggca gactggtcaa acaaatggat 180
tttacagagg cttacgcgga cacgtgctct acagttggac ttgctgccag ggaaggcaat 240
gttaaagtct taaggaaact gctcaaaaag ggccgaagtg tcgatgttgc tgataacagg 300
ggatggatgc caattcatga agcagcttat cacaactctg tagaatgttt gcaaatgtta 360
attaatgcag attcatctga aaactacatt aagatgaaga cctttgaagg tttctgtgct 420
ttgcatctcg ctgcaagtca aggacattgg aaaatcgtac agattctttt agaagctggg 480
gcagatccta atgcaactac tttagaagaa acgacaccat tgtttttagc tgttgaaaat 540
ggacagatag atgtgttaag gctgttgctt caacacggag caaatgttaa tggatcccat 600
tctatgtgtg gatggaactc cttgcaccag gcttcttttc aggaaaatgc tgagatcata 660
aaattgcttc ttagaaaagg agcaaacaag gaatgccagg atgactttgg aatcacacct 720
ttatttgtgg ctgctcagta tggcaagcta gaaagcttga gcatacttat ttcatcaggt 780
gcaaatgtca attgtcaagc cttggacaaa gctacaccct tgttcattgc tgctcaagag 840
ggacacacaa aatgtgtgga gcttttgctc tccagtgggg cagatcctga tctttactgt 900
aatgaggaca gttggcagtt acctattcat gcagctgcac aaatgggcca tacaaaaatc 960
ttggacttgt taataccact tactaaccgg gcctgtgaca ctgggctaaa caaagtaagc 1020
cctgtttact cagcagtgtt tgggggacat gaagattgcc tagaaatatt actccggaat 1080
ggctacagcc cagacgccca ggcgtgcctt gtttttggat tcagttctcc tgtgtgcatg 1140
gctttccaaa aggactgtga gttctttgga attgtgaaca ttcttttgaa atatggagcc 1200
cagataaatg aacttcattt ggcatactgc ctgaagtacg agaagttttc gatatttcgc 1260
tactttttga ggaaaggttg ctcattggga ccatggaacc atatatatga atttgtaaat 1320
catgcaatta aagcacaagc aaaatataag gagtggttgc cacatcttct ggttgctgga 1380
tttgacccac tgattctact gtgcaattct tggattgact cagtcagcat tgacacccttt 1440
atcttcactt tggagtttac taattggaag acacttgcac cagctgttga aaggatgctc 1500
tctgctcgtg cctcaaacgc ttggattcta cagcaacata ttgccactgt tccatccctg 1560
acccatcttt gtcgtttgga aattcggtcc agtctaaaat cagaacgtct acggtctgac 1620
agttatatta gtcagctgcc acttcccaga agcctacata attatttgct ctatgaagac 1680
gttctgagga tgtatgaagt tccagaactg gcagctattc aagatggata aatcagtgaa 1740
actacttaac acagctaatt tttttctctg aaaaatcatc gagacaaaag agccacagag 1800
tacaagtttt tatgatttta tagtcaaaag atgattattg attgtcagat aggttaggtt 1860
ttggggggcc agtagttcag tgagaatgtt tatgtttaca actagccttc ccagtaaaaa 1920
aaaaaaaaaa aaaaaaaaaa aattgtaaac atcacttata ttactttatt gcagcttcat 1980

```
caccagtaca ttatatgttg taatatttat ttacctgatc attttgatca ttttctgctt 2040
tattttgcta ataaactgtg atgttacttc tagtgctaaa catggcatat ttccacctat 2100
gattcgtgtt tacctggtat taggagctca gaatggaatg cataaagctt cactggaagt 2160
gtatacaact gtggtgtaga atctgttatt attatcatta ttattttatt tagacttgac 2220
tatctcttat gtttattaaa gaacatgttt tcctaaaaaa aaa                   2263
```